

Automatic Identification Of Cross-Document Structural Relationships

Abstract

Analysis on inter-document relationship is one of the important studies in multi document analysis. In this paper, we will focus on some special properties that multi document articles hold, specifically news articles. Information across news articles reporting on the same story are often related. Cross-document Structure Theory (CST) gives the relationship between pairs of sentences from different documents. For example, two sentences might have relationships such as identical, overlapping or contradicting. Our aim here is to automatically identify some of these CST relationships. We applied the well known machine learning technique, SVMs for this purpose and obtained some comparable results.